#2 OIPE

RAW SEQUENCE LISTING DATE: 04/09/2002
PATENT APPLICATION: US/09/924,841 TIME: 11:25:27

Input Set : N:\Crf3\RULE60\09924841.raw
Output Set: N:\CRF3\04092002\I924841.raw

<110> APPLICANT: Dilley, David R
Kadyrzhanova, Dina K
Wang, Zhenyong
Warner, Toni M
<120> TITLE OF INVENTION: Modified Synthetases To Produce Penicillins and Cephalosporins Under the Control of Bicarbonate
<130> FILE REFERENCE: MSU41-453
<140> CURRENT APPLICATION NUMBER: 09/924,841
<141> CURRENT FILING DATE: 2001-08-08
<150> PRIOR APPLICATION NUMBER: US/09/413,231
<151> PRIOR FILING DATE: 1999-10-06
<160> NUMBER OF SEQ ID NOS: 18
<170> SOFTWARE: PatentIn Ver. 2.0
<210> SEQ ID NO: 1
<211> LENGTH: 331
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: modified IPNS from Emericella nidulans strain b1A1
<221> NAME/KEY: MUTAGEN
<222> LOCATION: (212)
<223> OTHER INFORMATION: Glu212 in native IPNS modified to Arg
<400> SEQUENCE: 1

```
Met Gly Ser Val Ser Lys Ala Asn Val Pro Lys Ile Asp Val Ser Pro
 1               5                  10                  15
Leu Phe Gly Asp Asp Gln Ala Ala Lys Met Arg Val Ala Gln Gln Ile
            20                  25                  30
Asp Ala Ala Ser Arg Asp Thr Gly Phe Phe Tyr Ala Val Asn His Gly
        35                  40                  45
Ile Asn Val Gln Arg Leu Ser Gln Lys Thr Lys Glu Phe His Met Ser
    50                  55                  60
Ile Thr Pro Glu Glu Lys Trp Asp Leu Ala Ile Arg Ala Tyr Asn Lys
65                  70                  75                  80
Glu His Gln Asp Gln Val Arg Ala Gly Tyr Tyr Leu Ser Ile Pro Gly
                85                  90                  95
Lys Lys Ala Val Glu Ser Phe Cys Tyr Leu Asn Pro Asn Phe Thr Pro
            100                 105                 110
Asp His Pro Arg Ile Gln Ala Lys Thr Pro Thr His Glu Val Asn Val
        115                 120                 125
Trp Pro Asp Glu Thr Lys His Pro Gly Phe Gln Asp Phe Ala Glu Gln
    130                 135                 140
Tyr Tyr Trp Asp Val Phe Gly Leu Ser Ser Ala Leu Leu Lys Gly Tyr
```

RAW SEQUENCE LISTING DATE: 04/09/2002
PATENT APPLICATION: US/09/924,841 TIME: 11:25:27

Input Set : N:\Crf3\RULE60\0924841.raw
Output Set: N:\CRF3\04092002\I924841.raw

```
48       145                 150                 155                 160
49       Ala Leu Ala Leu Gly Lys Glu Glu Asn Phe Phe Ala Arg His Phe Lys
50                       165                 170                 175
51       Pro Asp Asp Thr Leu Ala Ser Val Val Leu Ile Arg Tyr Pro Tyr Leu
52                   180                 185                 190
53       Asp Pro Tyr Pro Glu Ala Ala Ile Lys Thr Ala Ala Asp Gly Thr Lys
54               195                 200                 205
55       Leu Ser Phe Arg Trp His Glu Asp Val Ser Leu Ile Thr Val Leu Tyr
56           210                 215                 220
57       Gln Ser Asn Val Gln Asn Leu Gln Val Glu Thr Ala Ala Gly Tyr Gln
58       225                 230                 235                 240
59       Asp Ile Glu Ala Asp Asp Thr Gly Tyr Leu Ile Asn Cys Gly Ser Tyr
60                       245                 250                 255
61       Met Ala His Leu Thr Asn Asn Tyr Tyr Lys Ala Pro Ile His Arg Val
62                   260                 265                 270
63       Lys Trp Val Asn Ala Glu Arg Gln Ser Leu Pro Phe Phe Val Asn Leu
64               275                 280                 285
65       Gly Tyr Asp Ser Val Ile Asp Pro Phe Asp Pro Arg Glu Pro Asn Gly
66           290                 295                 300
67       Lys Ser Asp Arg Glu Pro Leu Ser Tyr Gly Asp Tyr Leu Gln Asn Gly
68       305                 310                 315                 320
69       Leu Val Ser Leu Ile Asn Lys Asn Gly Gln Thr
70                       325                 330
72 <210> SEQ ID NO: 2
73 <211> LENGTH: 331
74 <212> TYPE: PRT
75 <213> ORGANISM: Artificial Sequence
76 <220> FEATURE:
77 <223> OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
78       from Emericella nidulans strain FGSC-4
79 <221> NAME/KEY: MUTAGEN
80 <222> LOCATION: (212)
81 <223> OTHER INFORMATION: Glu212 in native IPNS modified to Arg
82 <400> SEQUENCE: 2
83       Met Gly Ser Val Ser Lys Ala Asn Val Pro Lys Ile Asp Val Ser Pro
84         1               5                  10                  15
85       Leu Phe Gly Asp Asp Gln Ala Ala Lys Met Arg Val Ala Gln Gln Ile
86                    20                  25                  30
87       Asp Ala Ala Ser Arg Asp Thr Gly Phe Phe Tyr Ala Val Asn His Gly
88                35                  40                  45
89       Ile Asn Val Gln Arg Leu Ser Gln Lys Thr Lys Glu Phe His Met Ser
90            50                  55                  60
91       Ile Thr Pro Glu Glu Lys Trp Asp Leu Ala Ile Arg Ala Tyr Asn Lys
92        65                  70                  75                  80
93       Glu His Gln Asp Gln Val Arg Ala Gly Tyr Tyr Leu Ser Ile Pro Gly
94                        85                  90                  95
95       Lys Lys Ala Val Glu Ser Phe Cys Tyr Leu Asn Pro Asn Phe Thr Pro
96                    100                 105                 110
97       Asp His Pro Arg Ile Gln Ala Lys Thr Pro Thr His Glu Val Asn Val
```

RAW SEQUENCE LISTING DATE: 04/09/2002
PATENT APPLICATION: US/09/924,841 TIME: 11:25:27

Input Set : N:\Crf3\RULE60\09924841.raw
Output Set: N:\CRF3\04092002\I924841.raw

```
98                115                 120                 125
99        Trp Pro Asp Glu Thr Lys His Pro Gly Phe Gln Asp Phe Ala Glu Gln
100           130                 135                 140
101        Tyr Tyr Trp Asp Val Phe Gly Leu Ser Ser Ala Leu Leu Lys Gly Tyr
102        145                 150                 155                 160
103        Ala Leu Ala Leu Gly Lys Glu Glu Asn Phe Phe Ala Arg His Phe Lys
104                        165                 170                 175
105        Pro Asp Asp Thr Leu Ala Ser Val Val Leu Ile Arg Tyr Pro Tyr Leu
106                    180                 185                 190
107        Asp Pro Tyr Pro Glu Ala Ala Ile Lys Thr Ala Ala Asp Gly Thr Lys
108                195                 200                 205
109        Leu Ser Phe Arg Trp His Glu Asp Val Ser Leu Ile Thr Val Leu Tyr
110            210                 215                 220
111        Gln Ser Asn Val Gln Asn Leu Gln Val Glu Thr Ala Ala Gly Tyr Gln
112        225                 230                 235                 240
113        Asp Ile Glu Ala Asp Asp Thr Gly Tyr Leu Ile Asn Cys Gly Ser Tyr
114                        245                 250                 255
115        Met Ala His Leu Thr Asn Asn Tyr Tyr Lys Ala Pro Ile His Arg Val
116                    260                 265                 270
117        Lys Trp Val Asn Ala Glu Arg Gln Ser Leu Pro Phe Phe Val Asn Leu
118                275                 280                 285
119        Gly Tyr Asp Ser Val Ile Asp Pro Phe Asp Pro Arg Glu Pro Asn Gly
120            290                 295                 300
121        Lys Ser Asp Arg Glu Pro Leu Ser Tyr Gly Asp Tyr Leu Gln Asn Gly
122        305                 310                 315                 320
123        Leu Val Ser Leu Ile Asn Lys Asn Gly Gln Thr
124                        325                 330
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 331
128 <212> TYPE: PRT
129 <213> ORGANISM: Artificial Sequence
130 <220> FEATURE:
131 <223> OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
132       from Emericella nidulans strain bioA1
133 <221> NAME/KEY: MUTAGEN
134 <222> LOCATION: (212)
135 <223> OTHER INFORMATION: Glu212 in native IPNS modified to Arg
136 <400> SEQUENCE: 3
137        Met Gly Ser Val Ser Lys Ala Asn Val Pro Lys Ile Asp Val Ser Pro
138          1               5                  10                  15
139        Leu Phe Gly Asp Asp Gln Ala Ala Lys Met Arg Val Ala Gln Gln Ile
140                     20                  25                  30
141        Asp Ala Ala Ser Arg Asp Thr Gly Phe Phe Tyr Ala Val Asn His Gly
142                 35                  40                  45
143        Ile Asn Val Gln Arg Leu Ser Gln Lys Thr Lys Glu Phe His Met Ser
144             50                  55                  60
145        Ile Thr Pro Glu Glu Lys Trp Asp Leu Ala Ile Arg Ala Tyr Asn Lys
146         65                  70                  75                  80
147        Glu His Gln Asp Gln Val Arg Ala Gly Tyr Tyr Leu Ser Ile Pro Gly
```

RAW SEQUENCE LISTING DATE: 04/09/2002
PATENT APPLICATION: US/09/924,841 TIME: 11:25:27

Input Set : N:\Crf3\RULE60\09924841.raw
Output Set: N:\CRF3\04092002\I924841.raw

```
148                         85                  90                  95
149       Lys Lys Ala Val Glu Ser Phe Cys Tyr Leu Asn Pro Asn Phe Thr Pro
150                     100                 105                 110
151       Asp His Pro Arg Ile Gln Ala Lys Thr Pro Thr His Glu Val Asn Val
152                 115                 120                 125
153       Trp Pro Asp Glu Thr Lys His Pro Gly Phe Gln Asp Phe Ala Glu Gln
154             130                 135                 140
155       Tyr Tyr Trp Asp Val Phe Gly Leu Ser Ser Ala Leu Leu Lys Gly Tyr
156       145                 150                 155                 160
157       Ala Leu Ala Leu Gly Lys Glu Glu Asn Phe Phe Ala Arg His Phe Lys
158                         165                 170                 175
159       Pro Asp Asp Thr Leu Ala Ser Val Val Leu Ile Arg Tyr Pro Tyr Leu
160                     180                 185                 190
161       Asp Pro Tyr Pro Glu Ala Ala Ile Lys Thr Ala Ala Asp Gly Thr Lys
162                 195                 200                 205
163       Leu Ser Phe Arg Trp His Glu Asp Val Ser Leu Ile Thr Val Leu Tyr
164             210                 215                 220
165       Gln Ser Asn Val Gln Asn Leu Gln Val Glu Thr Ala Ala Gly Tyr Gln
166       225                 230                 235                 240
167       Asp Ile Glu Ala Asp Asp Thr Gly Tyr Leu Ile Asn Cys Gly Ser Tyr
168                         245                 250                 255
169       Met Ala His Leu Thr Asn Asn Tyr Tyr Lys Ala Pro Ile His Arg Val
170                     260                 265                 270
171       Lys Trp Val Asn Ala Glu Arg Gln Ser Leu Pro Phe Phe Val Asn Leu
172                 275                 280                 285
173       Gly Tyr Asp Ser Val Ile Asp Pro Phe Asp Pro Arg Glu Pro Asn Gly
174             290                 295                 300
175       Lys Ser Asp Arg Glu Pro Leu Ser Tyr Gly Asp Tyr Leu Gln Asn Gly
176       305                 310                 315                 320
177       Leu Val Ser Leu Ile Asn Lys Asn Gly Gln Thr
178                         325                 330
180 <210> SEQ ID NO: 4
181 <211> LENGTH: 331
182 <212> TYPE: PRT
183 <213> ORGANISM: Artificial Sequence
184 <220> FEATURE:
185 <223> OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
186       from Penicillium chrysogenum
187 <221> NAME/KEY: MUTAGEN
188 <222> LOCATION: (212)
189 <223> OTHER INFORMATION: Glu212 in native IPNS modified to Arg
190 <400> SEQUENCE: 4
191       Met Ala Ser Thr Pro Lys Ala Asn Val Pro Lys Ile Asp Val Ser Pro
192         1               5                  10                  15
193       Leu Phe Gly Asp Asn Met Glu Glu Lys Met Lys Val Ala Arg Ala Ile
194                     20                  25                  30
195       Asp Ala Ala Ser Arg Asp Thr Gly Phe Phe Tyr Ala Val Asn His Gly
196                 35                  40                  45
197       Val Asp Val Lys Arg Leu Ser Asn Lys Thr Arg Glu Phe His Phe Ser
```

RAW SEQUENCE LISTING DATE: 04/09/2002
PATENT APPLICATION: US/09/924,841 TIME: 11:25:27

Input Set : N:\Crf3\RULE60\09924841.raw
Output Set: N:\CRF3\04092002\I924841.raw

```
198            50                  55                  60
199       Ile Thr Asp Glu Glu Lys Trp Asp Leu Ala Ile Arg Ala Tyr Asn Lys
200        65                  70                  75                  80
201       Glu His Gln Asp Gln Ile Arg Ala Gly Tyr Tyr Leu Ser Ile Pro Glu
202                        85                  90                  95
203       Lys Lys Ala Val Glu Ser Phe Cys Tyr Leu Asn Pro Asn Phe Lys Pro
204                   100                 105                 110
205       Asp His Pro Leu Ile Gln Ser Lys Thr Pro Thr His Glu Val Asn Val
206               115                 120                 125
207       Trp Pro Asp Glu Lys Lys His Pro Gly Phe Arg Glu Phe Ala Glu Gln
208           130                 135                 140
209       Tyr Tyr Trp Asp Val Phe Gly Leu Ser Ser Ala Leu Leu Arg Gly Tyr
210       145                 150                 155                 160
211       Ala Leu Ala Leu Gly Lys Glu Glu Asp Phe Phe Ser Arg His Phe Lys
212                        165                 170                 175
213       Lys Glu Asp Ala Leu Ser Ser Val Val Leu Ile Arg Tyr Pro Tyr Leu
214                   180                 185                 190
215       Asn Pro Ile Pro Pro Ala Ala Ile Lys Thr Ala Glu Asp Gly Thr Lys
216               195                 200                 205
217       Leu Ser Phe Arg Trp His Glu Asp Val Ser Leu Ile Thr Val Leu Tyr
218           210                 215                 220
219       Gln Ser Asp Val Ala Asn Leu Gln Val Glu Met Pro Gln Gly Tyr Leu
220       225                 230                 235                 240
221       Asp Ile Glu Ala Asp Asp Asn Ala Tyr Leu Val Asn Cys Gly Ser Tyr
222                        245                 250                 255
223       Met Ala His Ile Thr Asn Asn Tyr Tyr Pro Ala Pro Ile His Arg Val
224                   260                 265                 270
225       Lys Trp Val Asn Glu Glu Arg Gln Ser Leu Pro Phe Phe Val Asn Leu
226               275                 280                 285
227       Gly Phe Asn Asp Thr Val Gln Pro Trp Asp Pro Ser Lys Glu Asp Gly
228           290                 295                 300
229       Lys Thr Asp Gln Arg Pro Ile Ser Tyr Gly Asp Tyr Leu Gln Asn Gly
230       305                 310                 315                 320
231       Leu Val Ser Leu Ile Asn Lys Asn Gly Gln Thr
232                        325                 330
234 <210> SEQ ID NO: 5
235 <211> LENGTH: 329
236 <212> TYPE: PRT
237 <213> ORGANISM: Artificial Sequence
238 <220> FEATURE:
239 <223> OTHER INFORMATION: Description of Artificial Sequence: modified IPNS
240       from Streptomyces clavuligerus
241 <221> NAME/KEY: MUTAGEN
242 <222> LOCATION: (210)
243 <223> OTHER INFORMATION: Glu210 in native IPNS modified to Arg
244 <400> SEQUENCE: 5
245       Met Pro Val Leu Met Pro Ser Ala His Val Pro Thr Ile Asp Ile Ser
246         1               5                  10                  15
247       Pro Leu Phe Gly Thr Asp Ala Ala Ala Lys Lys Arg Val Ala Glu Glu
```

VERIFICATION SUMMARY
PATENT APPLICATION: US/09/924,841

DATE: 04/09/2002
TIME: 11:25:28

Input Set : N:\Crf3\RULE60\09924841.raw
Output Set: N:\CRF3\04092002\I924841.raw